

**Figure 1A**

**NOVTRAN cDNA sequence:**

atgcagtggtcctgtctggcctgtaccctcctcagggtcctcccacatgtctgtctcctgagagaccctgtgcctgtgcccacagggaccaagctcttc  
cactcctgtatcacctcaacgaacccatgcgcctcctcctggaggttgctgttgagctgcaggcatcacccctggactgtagggtctgagcaccggcc  
ctgtccatatccatccctgcatgcctctccgttcaccgactcctcaacagaccatccctgtcctctcaacaggccccgctctgtctggggaacca  
cggacagaggccttccatccccaggcctgaaggccagagtaggtgggaccatcctcgcgaagccggcctcaattctcaaggccatgccgtggagc  
cagtgccatctggaccctctgggtcaagcaaagggtgtgtgctaatcaaaggcaggccctcaggatgccaaaggccccgcaatgccagtggaccg  
tgaaaaccttctgtgacaaacctgcagtgccttctgtctccagctgctctccagctcctcatcaaggtgaaacagagcaggagcgagtaatg  
cggaaattgacttgcaaagtcgggcccgtcgggattacaattcaaggctgctgctgaaactcgggcagatcccagctgcaaagggcagttccttcctga  
gctgcagaacgtgtctggaggggttggtcagcccagggtcccaggaaccactgcaaggtggggcgggccctcagagcccttccagagctggg  
ggctggtagccccctttggcttggagaaggtagtaccacaacccatccccagggccgactgcggaagggtgtgactggccccctgtgtctctgtg  
gaccagtgtccactgtgactctcccaggccagccgaacctggcacacactgggtgtccctaataagccatggagggtattgtggcatggagagctgtc  
gattccagaaacctcctggacatagggctgggagctcatctgcagaagctgcctga

**Figure 1B**

**NOVTRAN Protein Sequence:**

MQWSCLACTLLRVLPVLSLLRDPVPVPTGTKLFHSCITSTNPCASFLEVAVEAAGITPW  
TVGSEHPPCPYPSLHASPF TDSFNRPSPAPLNRPRSAGEPRTEAFPSPGLKARVGGTILAE  
AGLNSQGHAVEPVPSGPSGSSKGCVLIKGRPSRMPKARECPVDRENLLL TNPAVPSLLQL  
LSSSPCIKVETE QERSNAEFDLQSRARDYNSRLL LKLQIPAAKGSSFLELQNVSGGVG  
SARGPRNHCKVGAGPQSPFPELGAGSPPLALEKVSTQPIQARLRKGVDWPPVSPGDQC  
PLCTLPGQP NLAHTGCSLNSHGGYCGMESCRFQKPPGHRAGSSSAEAA

**Figure 2A**

**BlastN for NOVTRAN:**

gb:GENBANK-ID:H55724|acc:H55724 CHR220663 Chromosome 22 exon Homo sapiens  
genomic clone C22\_916 5', mRNA sequence - Homo sapiens, 171 bp.

Minus Strand HSPs:

Score = 585 (87.8 bits), Expect = 1.0e-19, P = 1.0e-19  
Identities = 117/117 (100%), Positives = 117/117 (100%), Strand = Minus /  
Plus

```
Query:   735 ACCTCGGGCTGAGCCAACCCCTCCAGACACGTTCTGCAGCTCGAGGAAGGAAGTGCCTT 676
          |||
Sbjct:   18 ACCTCGGGCTGAGCCAACCCCTCCAGACACGTTCTGCAGCTCGAGGAAGGAAGTGCCTT 77

Query:   675 TGCAGCTGGGATCTGCCCGAGTTTCAGCAGCAGCCTTGAATTGTAATCCCGAGCGGC 619
          |||
Sbjct:   78 TGCAGCTGGGATCTGCCCGAGTTTCAGCAGCAGCCTTGAATTGTAATCCCGAGCGGC 134
```

gb:GENBANK-ID:H55724|acc:H55724 CHR220663 Chromosome 22 exon Homo sapiens  
genomic clone C22\_916 5', mRNA sequence - Homo sapiens, 171 bp.

Figure 2B

**BlastX for NOVTRAN:**

No Hits:

ptnr:SWISSPROT-ACC:P22357 ANThER-SPECIFIC PROTEIN SF18...	-3	102	0.0093	1
ptnr:SPTREMBL-ACC:Q26876 PROTECTIVE RECOMBINANT ANTIGE...	-3	87	0.018	1
ptnr:TREMBLNEW-ACC:AAD41775 PRO-ALPHA-2(I) COLLAGEN - ...	-3	86	0.061	2
ptnr:PIR-ID:S46272 anther-specific protein - common su...	-3	94	0.073	1
ptnr:SPTREMBL-ACC:Q26874 PROTECTIVE RECOMBINANT ANTIGE...	-3	78	0.16	1
ptnr:SPTREMBL-ACC:Q92029 ALPHA-1 COLLAGEN TYPE III - G...	-3	86	0.19	3
ptnr:SWISSPROT-ACC:P46804 SPIDROIN 2 (DRAGLINE SILK FI...	-3	101	0.24	1
ptnr:SPTREMBL-ACC:O57808 103AA LONG HYPOTHETICAL PROTE...	+2	76	0.25	1
ptnr:SWISSNEW-ACC:P41484 PROLINE-RICH ANTIGEN (36 KD A...	-3	95	0.26	1

**CuraBLASTX Analysis of AC007663\_A**

PUBLIC DATABASE

Query= AC007663\_A Cura\_109 transmembrane protein  
(1047 letters)

Database: /opt/database/public/blast/protnr  
577,633 sequences; 178,813,065 total letters.

			Smallest Sum	Reading Frame	High Score	Probability P(N)	N
Sequences producing High-scoring Segment Pairs:							
ptnr:SWISSPROT-ACC:P22357 ANThER-SPECIFIC PROTEIN SF18...	-3	102	0.015	1			
ptnr:pir-id:S12246 anther-specific protein SF18 precur...	-3	102	0.015	1			

PATENT DATABASE

**CuraBLASTX Analysis of AC007663\_A**

Query= AC007663\_A Cura\_109 transmembrane protein  
(1047 letters) of query sequence in all 6 reading frames

Database: /opt/database/licensed/blast/geneseq\_aa  
349,121 sequences; 51,277,408 total letters.

			Smallest Sum	Reading Frame	High Score	Probability P(N)	N
Sequences producing High-scoring Segment Pairs:							
patp:R28916 Type III procollagen (prior art) - Homo sa...	+1	81	0.059	3			
patp:W03454 Mutant mouse c-AKT non-translation region ...	+1	81	0.075	1			
patp:G04146 Arabidopsis thaliana protein fragment SEQ ...	-3	80	0.095	1			

### Figure 3A

#### NOVNEUR cDNA sequence:

Regions 135319 to 135129 (1-190bp) and 134487 to 134314 (191 to 364bp) of the genomic clone (in - frame). Further extended using ESTs AI739078, AI767488, and AI796217.

AGCGCGCCCGAACGAAGCCGCGGCCCGGGCACAGCATGGCCCGCGGCGGGAGGGCGCTCGGATGTTTCGGC  
AGCCTCCTGCACTTCGCCCTGCTCGCTGCCGGCGTCGTCCCGCTCAGCTGGGATCTCCCGGAGCCCCGCA  
GCCGAGCCAGCAAGATCCGAGTGCCTCGCGAGGCAAGCTCTGGGCCATCGGTCACTTCATGGGCAAGAA  
GAGTCTGGAGCCTTCCAGCCCATCCCCATTGGGGACAGCTCCCCACACCTCCCTGAGGGACCAGCGACTG  
CAGCTGAGTCATGATCTGCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCCG  
CACCCCAAATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAATGGGGCAGACA  
CAACAGCGTGGCTTAGATTGTGCCCACCCAGGGAAGGTGCTGAATGGGACCTGTTGATGGCCCCATCTG  
GATGTAAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATTTCTGGCTGGGTCAACCAGAAATATC  
GCTGATGCAGACACAGATTATGTTCTGCTGTATTTCCTGCTTCCCTGTTGAATTGGTGAATAAACCTT  
GCTCTATACATACAAA

### Figure 3B

#### NOVNEUR Protein sequence:

MFGSLHLHFALLAAGVVPLSWDLPEPRSRASKIRVHSRGKLWAIGHFMGKKSLEPSSPSPLGTAPHTSLRDQRLQLSH  
DLLGILLKKALGVSLSRPAPQIQYRLLLVQILQK

NOVNEUR cDNA sequence: AGCGCGCCCGAACGAAGCCGCGGCCCGGGCACAGCATGGCCCGCGGCGGGAGGGCGCTCGGATGTTTCGGC  
AGCCTCCTGCACTTCGCCCTGCTCGCTGCCGGCGTCGTCCCGCTCAGCTGGGATCTCCCGGAGCCCCGCA  
GCCGAGCCAGCAAGATCCGAGTGCCTCGCGAGGCAAGCTCTGGGCCATCGGTCACTTCATGGGCAAGAA  
GAGTCTGGAGCCTTCCAGCCCATCCCCATTGGGGACAGCTCCCCACACCTCCCTGAGGGACCAGCGACTG  
CAGCTGAGTCATGATCTGCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCCG  
CACCCCAAATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAATGGGGCAGACA  
CAACAGCGTGGCTTAGATTGTGCCCACCCAGGGAAGGTGCTGAATGGGACCTGTTGATGGCCCCATCTG  
GATGTAAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATTTCTGGCTGGGTCAACCAGAAATATC  
GCTGATGCAGACACAGATTATGTTCTGCTGTATTTCCTGCTTCCCTGTTGAATTGGTGAATAAACCTT  
GCTCTATACATACAAA

Figure 4A

**BlastN for NOVNEUR:**

gb:GENBANK-ID:HUMNKB|acc:M21551 Human neuromedin B mRNA, complete cds -  
Homo sapiens, 640 bp (RNA).

Top Previous Match Next Match  
Length = 640

Plus Strand HSPs:

Score = 3055 (458.4 bits), Expect = 9.1e-132, P = 9.1e-132

Identities = 629/639 (98%), Positives = 629/639 (98%), Strand = Plus / Plus

Query: 2 GCGCGCCCGAACGAAGCCGCGGCCCGGGCACAGC-ATGGCCCG-CGG-CGGGAGGGCGCT 58  
GCGCGCCCGAACGAAGCCGCGGCCCGGGCACAGC ATGGCCCG CGG CGGG GGGCGCT  
Sbjct: 2 GCGCGCCCGAACGAAGCCGCGGCCCGGGCACAGCCATGGCCCGCGGGCGGG-GGGCGCT 60

Query: 59 CGGATGTTTCGGCAGCCTCCTGCACTTCGCCCTGCTCGCTGCCGGCGTCGTCCCGCTCAGC 118  
CGGATGTTTCGGCAGCCTCCTGC CTCGCCCTGCTCGCTGCCGGCGTCG CCCGCTCAGC  
Sbjct: 61 CGGATGTTTCGGCAGCCTCCTGCTCTTCGCCCTGCTCGCTGCCGGCGTCGCCCCGCTCAGC 120

Query: 119 TGGGATCTCCCGAGCCCCGCAGCCGAGCCAGCAAGATCCGAGTGCACTCGCGAGGCAAG 178  
TGGGATCTCCCGAGCCCCGCAGCCGAGCCAGCAAGATCCGAGTGCACTCGCGAGGCAA  
Sbjct: 121 TGGGATCTCCCGAGCCCCGCAGCCGAGCCAGCAAGATCCGAGTGCACTCGCGAGGCAAC 180

Query: 179 CTCTGGGGCCATCGGTCACTTCATGGGCAAGAAGAGTCTGGAGCCTTCCAGCCCATCCCCA 238  
CTCTGGGGCCA CGGTCACTTCATGGGCAAGAAGAGTCTGGAGCCTTCCAGCCCATCCC A  
Sbjct: 181 CTCTGGGGCCACCGGTCACTTCATGGGCAAGAAGAGTCTGGAGCCTTCCAGCCCATCCC-A 239

Query: 239 TTGGGGACAGCTCCCCACACCTCCC-TGAGGGACCAGCGACTGCAGCTGAGTCATGATCT 297  
TTGGGGACAGCTCCCCACACCTCCC TGAGGGACCAGCGACTGCAGCTGAGTCATGATCT  
Sbjct: 240 TTGGGGACAGCTCCCCACACCTCCCCCTGAGGGACCAGCGACTGCAGCTGAGTCATGATCT 299

Query: 298 GCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCCCGCACCCCA 357  
GCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCCCGCACCCCA  
Sbjct: 300 GCTCGGAATCCTCCTGCTAAAGAAGGCTCTGGGCGTGAGCCTCAGCCGCCCCCGCACCCCA 359

Query: 358 AATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAAT-GGGGCA 416  
AATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAAT GGGGCA  
Sbjct: 360 AATCCAGTACAGGAGGCTGCTGGTACAAATACTGCAGAAATGACACCAATAATAGGGGCA 419

Query: 417 GACACAACAGCGTGGCTTAGATTGTGCCACCCAGGGAAGGTGCTGAATGGGACCCTGTT 476  
GACACAACAGCGTGGCTTAGATTGTGCCACCCAGGGAAGGTGCTGAATGGGACCCTGTT  
Sbjct: 420 GACACAACAGCGTGGCTTAGATTGTGCCACCCAGGGAAGGTGCTGAATGGGACCCTGTT 479

Query: 477 GATGGCCCCATCTGGATGTAAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATT 536  
GATGGCCCCATCTGGATGTAAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATT  
Sbjct: 480 GATGGCCCCATCTGGATGTAAATCCTGAGCTCAAATCTCTGTTACTCCATTACTGTGATT 539

Query: 537 TCTGGCTGGGTCACCAGAAATATCGCTGATGCAGACACAGATTATGTTCTCTGCTGTATTT 596  
TCTGGCTGGGTCACCAGAAATATCGCTGATGCAGACACAGATTATGTTCTCTGCTGTATTT  
Sbjct: 540 TCTGGCTGGGTCACCAGAAATATCGCTGATGCAGACACAGATTATGTTCTCTGCTGTATTT 599

Query: 597 CCTGCTTCCCTGTTGAATTGGTGAATAAAACCTTGCTCTAT 637  
CCTGCTTCCCTGTTGAATTGGTGAATAAAACCTTGCTCT T  
Sbjct: 600 CCTGCTTCCCTGTTGAATTGGTGAATAAAACCTTGCTCTTT 640

## Figure 4B

### BlastX for NOVNEUR:

ptnr:SWISSPROT-ACC:P08949 NEUROMEDIN B-32 PRECURSOR [CONTAINS: NEUROMEDIN  
B] - Homo sapiens (Human), 121 aa.

Top Previous Match Next Match

Length = 121

Plus Strand HSPs:

Score = 513 (180.6 bits), Expect = 1.9e-48, P = 1.9e-48

Identities = 105/118 (88%), Positives = 105/118 (88%), Frame = +2

Query: 44 RREGARMFGSLLHFALLAAGVVPLSWDLPEPRSRASKIRVHSRGKLWAIGHFMGKKSLEP 223

R GARMFGSLL FALLAAGV PLSWDLPEPRSRASKIRVHSRG LWA GHFMGKKSLEP

Sbjct: 4 RAGGARMFGSLLLLFALLAAGVAPLSWDLPEPRSRASKIRVHSRGNLWATGHFMGKKSLEP 63

Query: 224 SSPSPLGTAPHTSLRDQRLQLSHDLLGILLKKKALGVSLSRPAPQIQYRLLVQILQK 397

SSPS G P LRDQRLQLSHDLLGILLKKKALGVSLSRPAPQIQYRLLVQILQK

Sbjct: 64 SSPSHWGQLPTPLRDQRLQLSHDLLGILLKKKALGVSLSRPAPQIQYRLLVQILQK 121

Query: 44 RREGARMFGSLLHFALLAAGVVPLSWDLPEPRSRASKIRVHSRGKLWAIGHFMGKKSLEP 223  
R GARMFGSLL FALLAAGV PLSWDLPEPRSRASKIRVHSRG LWA GHFMGKKSLEP  
Sbjct: 4 RAGGARMFGSLLLLFALLAAGVAPLSWDLPEPRSRASKIRVHSRGNLWATGHFMGKKSLEP 63  
  
Query: 224 SSPSPLGTAPHTSLRDQRLQLSHDLLGILLKKKALGVSLSRPAPQIQYRLLVQILQK 397  
SSPS G P LRDQRLQLSHDLLGILLKKKALGVSLSRPAPQIQYRLLVQILQK  
Sbjct: 64 SSPSHWGQLPTPLRDQRLQLSHDLLGILLKKKALGVSLSRPAPQIQYRLLVQILQK 121

Figure 5

**ClustalW for NOVNEUR:**

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g., L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

**Sequences used:**

- 1) A37178 (neuromedin B precursor – rat): Locus A37178, accession A37178, PID - g112182
- 2) NEUB\_HUMAN (NEUROMEDIN B-32 PRECURSOR): Locus - NEUB\_HUMAN, accession - P08949, PID - g1346684

```
Neuromedin_New_REVCOMP  -----MFGSLLHFALLAAGVPLSWDLPEPRSRASKIRVHSRGK|LWAI|GHFMGKKS
A37178                   MTRQAGSTWLLRGLLLFALFVSGITPFSWDLPEPRSRASKIRVHPRGNLWATGHFMGKKS
NEUB_HUMAN               MARRAGGARMFGSLLLFALLAAGVPLSWDLPEPRSRASKIRVHSRGNLWATGHFMGKKS

Neuromedin_New_REVCOMP  LEPSSPSP|GTAPHTSLRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRRLVQILQ
A37178                   LEP|SLSLVGTAPPI|QRLQLSHDLLRI|LLQKALGMNLSGPAPPIQYRRLQK---
NEUB_HUMAN               LEPSSPSHWGQLPTPPLRDQRLQLSHDLLGILLKKALGVSLSRPAPQIQYRRLVQILQ

Neuromedin_New_REVCOMP  K
A37178                   -
NEUB_HUMAN               K
```

**Figure 6A**

**NOVGON cDNA sequence:**

1  
ATGAAGCTGGCATTCTCTTCTTGGCCCCATGGCCCTCCTCCTTCTGGCTGGCTATGGCTGTGTCTCGGTGCCTCCAG  
81  
TGGGAACCTGCGCACCTTTGTGGGCTGTGCCGTGAGGGAGTTTACTTTCTGGCCAAGAAGCCAGGCTGCAGGGGCCTTC  
161  
GGATCACACGGATGCCTGCTGGGGTCGCTGTGAGACCTGGGAGAAACCCATTCTGGAACCCCCCTATATTGAAGCCCAT  
241  
CATCGAGTCTGTACCTACAACGAGACCAAACAGGTGACTGTCAAGCTGCCCCAACTGTGCCCCGGGAGTCGACCCCTTCTA  
321  
CACCTATCCCGTGGCCATCCGCTGTGACTGCGGAGCCTGCTCCACTGCCACCACGGAGCTGAGGTTGATGCCAGGGGAAG  
401  
CTGCTGTGGCACTGGGCTTCTGGTGTGAGCGTAGGAGACAGGGATCTAGGACAACAGGGACCAGGTGGCGACATGCAGCT  
481  
GTAAGAGACAAGGTGAGTCTCCTGAAGGCAGTAGATGGTTGGAATGGGCTGCTTGGGGACCCAGCGAGCTCCCAGGGCCT  
561  
TTCTGCTTCTTCTGTACCCCTGTATTTCCCTTGGCTTTCCAAATTGACTCAGCTTCTGGTAAAGTTGGAACTTTTCCA  
641  
GCAAACAGACCTTCATCTTCTCCAGTGCAGAGATTACATTAGGAGGAACATGA

**Figure 6B**

**NOVGON Protein Sequence:**

MKLAFLFLGPMALLLLAGYGCVLGASSGNLRTFVGCAVREFTFLAKKPGCRGLRITTTDA  
CWGRCEWKEPPYIEAHRVCTYNETKQVTVKLPNCAPGVDPFYTYTPVAIRCDCG  
ACSTATTELRLMPGEAAVALGFWCQRRRQGSRTTGTRWRHA AVRDKVSLLKAVDGW  
NLLGDPASSQGLSASSCTPVFPLAFQIDSASGKVGNFSSKQTFIFSSAEITLGGT\*



## Figure 7A

### BlastN for NOVGON:

AF146151 Salmo salar gonadotropin II beta subunit mRNA,  
partial cds - Salmo salar, 266 bp (RNA).

Top Previous Match Next Match  
Length = 266

Plus Strand HSPs:

Score = 208 (31.2 bits), Expect = 0.043, P = 0.042

Identities = 124/205 (60%), Positives = 124/205 (60%), Strand = Plus / Plus

Query: 177 CTGCTGGGGTCGCTGTGAGACCTGGGAGAAACCCATTCTGGAACCCCCCTATATTGAAGC 236  
CTGC G GG C CTG G ACC GGAG CC TT T AA CCC AT TT A C  
Sbjct: 34 CTGCAGTGGCCACTGCGTCACCAAGGAG---CCGGTTTTC-AAGAGCCC-ATTTTCCACC 88

Query: 237 CCAT-C-ATCGAGTCTGTACCTACAACGA-GACCAAACAGGTGACTGTCAAGCTGCCCAA 293  
T C A C GT TG ACCTAC GA G CC A G AC TC CT CC A  
Sbjct: 89 GTGTACCAGCATGTGTGCACCTACCGGGACGTCCGCT-ATGAAACGATCCGCCTACCTGA 147

Query: 294 CTGTGCCCCGGGAGTCGACCC-CTTCTACACCTATCCCGTGGCCATCCGCTGTGACTGCG 352  
CTGT CCCC G GT GACC C T T CACCTA CC GTGGC T GCTGTGACTGC  
Sbjct: 148 CTGTCCCCCTTGGGTGGACCATCATGT-CACCTACCCTGTGGCTCTGAGCTGTGACTGCA 206

Query: 353 GAGCCTGCTCCACTGCCACCACGGAGCTG 381  
G CTG CA G CAC C GA CTG  
Sbjct: 207 GCCTCTGTAACATGGACACTTCTGA-CTG 234

Figure 7B

**BlastX for NOVGON:**

P01235 GONADOTROPIN BETA CHAIN PRECURSOR - *Cyprinus carpio* (Common carp), 144 aa.

Score = 203 (71.5 bits), Expect = 2.3e-16, P = 2.3e-16  
Identities = 37/85 (43%), Positives = 52/85 (61%)

Query: 42 TFLAKKPGC-RGLRITTDACWGRCEWKEPILEPPYIEAHHRVCTYNETKQVTVKLPNCA 100  
T +K GC + L + T C G C T E P+ + P+ + VCTY + + TV+LP+C  
Sbjct: 39 TVAVEKEGCPKCLVLQTTICSGHCLTKE-PVYKSPFSTVYQHVCTYRDVRYETVRLPDCP 97

Query: 101 PGVDPFYTYTPVAIRCDGACSTATTE 126  
PGVDP TYPVA+ CDC C+ T++  
Sbjct: 98 PGVDPHITYPVALSCDCSLCTMDTSD 123

>ptnr:SWISSNEW-ACC:Q9YGH2 GONADOTROPIN BETA-II CHAIN PRECURSOR (GTH-II-BETA)  
- *Clupea pallasii* (Pacific herring), 149 aa.

Plus Strand HSPs:

Score = 206 (72.5 bits), Expect = 1.0e-15, P = 1.0e-15  
Identities = 47/117 (40%), Positives = 63/117 (53%), Frame = +1

Query: 28 PMALLLLAGYGCVLG--ASSGNLRTFVGCAVREFTFLAKKPGC-RGLRITTDACWGRCEWKE 198  
P +LL CVL A NL+ C + T +K GC R L T C G C T  
Sbjct: 5 PECTILLLLCMCVLAVPAQCFNLQP---CVLVNETVSVEKEGCPRLVFRITICSGHCPT 61

Query: 199 WEKPILEPPYIEAHHRVCTYNETKQVTVKLPNCAPGVDPFYTYTPVAIRCDGACSTATTE 378  
E P+ + P+ + VCTY + T++LP+CA GVDP TYPVA+ C+C CS T++  
Sbjct: 62 KE-PVYKSPFSVNVQHVCYIGNFRYETIRLPDCADGVDPVLVTPVALSCDCSLCSMDTSD 120

>patp:R15106 hCG/bLH chimera, D10 - *Homo sapiens*, 145 aa.

Plus Strand HSPs:

Score = 188 (66.2 bits), Expect = 5.0e-14, P = 5.0e-14  
Identities = 40/101 (39%), Positives = 58/101 (57%), Frame = +1

Query: 76 SSGNLRTFVGCAVREFTFLAKKPGCRG-LRITTDACWGRCEWKE--PILEPPYIEAHHR 246  
S G LR C T A+K C + TT C G C + ++ P++ PP + R  
Sbjct: 1 SRGPLRPL--CQPINATLAAEKEACPVCITFTTSICAGYCPSMKRVLPVILPPMPQ---R 55

Query: 247 VCTYNETKQVTVKLPNCAPGVDPFYTYTPVAIRCDGACSTATTE 378  
VCTY+E + +V+LP C PGVDP ++PVA+ C CG C ++T+  
Sbjct: 56 VCTYHELRFASVRLPGCPPGVDPMVSPVALSCHCGPCRLSSTD 99

>patp:W65110 Equine chorionic gonadotropin beta-chain protein - *Equus caballus*, 169 aa.

Plus Strand HSPs:

Score = 175 (61.6 bits), Expect = 1.2e-12, P = 1.2e-12  
Identities = 43/116 (37%), Positives = 60/116 (51%), Frame = +1

Query: 31 MALLLLAGYGCVLGASSGNLRTFVGCAVREFTFLAKKPGCR-GLRITTDACWGRCEWKE 207  
+ L +L G V AS G LR C T A+K C + TT C G C + +  
Sbjct: 7 LLLWMLLSVGGVW-ASRGPLRPL--CRPINATLAAEKEACPICITFTTSICAGYCPSMVR 63

Query: 208 PILEPPYIEAHHR-VCTYNETKQVTVKLPNCAPGVDPFYTYTPVAIRCDGACSTATTE 378  
+ P + A + VCTY E + +++LP C PGVDP ++PVA+ C CG C TT+  
Sbjct: 64 VM--PAALPAIPQPVCTYRELRFASIRLPGCCPPGVDPMVSPVALSCHCGPCQIKTTD 119

## Figure 8

## ClustalW alignment for NOVGON:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g. L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

GTHB_CYPCA	MGTPVKI	L	V	V	R	N	H	I	L	F	S	V	V	L	L	A	V	A	Q	S	S	--	Y	L	P	P	C	E	P	V	N	E	T	V	A	V	E	K	E	G	C	P	K	C	L	V	L	Q	T	I			
Q98349_GTHB2_GOLDFISH	MGTPVKI	L	V	V	--	--	--	--	L	F	S	V	V	L	L	A	V	A	Q	S	S	--	Y	L	P	P	C	E	P	V	N	E	T	V	A	V	E	K	E	G	C	P	K	C	L	V	L	Q	T	I			
LSHB_BOVIN	MEMFOGI	L	I	W	--	--	--	--	L	L	G	V	A	G	V	W	A	S	R	G	P	--	L	R	E	L	C	Q	P	I	N	A	T	A	A	E	K	E	A	C	P	W	C	I	T	F	T	S	I				
LSHB_SHEEP	MEMLOGL	L	L	W	--	--	--	--	L	L	G	V	A	G	V	W	A	S	R	G	P	--	L	R	E	L	C	Q	P	I	N	A	T	A	A	E	K	E	A	C	P	W	C	I	T	F	T	S	I				
novel_gonadotropin	MKLAFLP	L	G	P	M	A	L	L	L	A	G	Y	G	V	L	G	S	S	G	N	L	R	T	F	V	G	C	A	V	R	E	T	F	L	A	K	K	P	G	C	R	G	-	R	T	T	D						
GTHB_CYPCA	CSGHCLT	-	K	E	P	V	K	S	P	F	S	T	V	Y	Q	H	V	C	T	Y	R	D	V	R	E	I	W	R	L	P	D	C	P	P	G	V	D	P	H	I	T	P	V	A	L	S	C	D	C	S	L		
Q98349_GTHB2_GOLDFISH	CSGHCLT	-	K	E	P	V	K	S	P	F	S	T	V	Y	Q	H	V	C	T	Y	R	D	V	R	E	I	W	R	L	P	D	C	P	P	G	V	D	P	H	I	T	P	V	A	L	S	C	D	C	S	L		
LSHB_BOVIN	CAGYCPS	-	M	K	R	V	L	P	V	I	L	P	P	M	P	Q	R	V	C	T	Y	H	E	L	R	E	A	S	W	R	L	P	G	C	P	P	G	V	D	P	M	V	S	P	P	V	A	L	S	C	H	G	P
LSHB_SHEEP	CAGYCPS	-	M	K	R	V	L	P	V	I	L	P	P	M	P	Q	R	V	C	T	Y	H	E	L	R	E	A	S	W	R	L	P	G	C	P	P	G	V	D	P	M	V	S	P	P	V	A	L	S	C	H	G	P
novel_gonadotropin	CWGRCE	T	W	E	K	P	I	L	E	P	P	Y	I	E	A	H	R	V	C	T	Y	N	E	I	K	V	I	W	K	L	P	N	C	A	R	G	V	D	P	F	Y	T	P	V	A	U	R	C	D	G	A		
GTHB_CYPCA	CTMDTSD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-				
Q98349_GTHB2_GOLDFISH	CTMDTSD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
LSHB_BOVIN	CR'LSSTD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
LSHB_SHEEP	CR'LSSTD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
novel_gonadotropin	C'ITATTE	L	R	L	M	P	G	E	A	A	V	A	L	G	F	W	C	Q	R	R	R	Q	S	R	T	G	T	R	W	R	H	A	A	R	D	K	Y	S	L	K	A	V	D	G	W	N	G	L	G				
GTHB_CYPCA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
Q98349_GTHB2_GOLDFISH	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
LSHB_BOVIN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
LSHB_SHEEP	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-					
novel_gonadotropin	D	P	A	S	S	Q	L	S	A	S	S	C	T	P	V	F	L	A	P	Q	I	D	S	A	S	G	K	V	G	N	F	S	S	K	Q	T	F	I	S	S	A	E	I	T	L	G	G	T					

## Figure 9A

### NOVINTRA A cDNA sequence:

Regions 116708 to 116812 (1-105bp), 117121 to 117248 (106-233p), and 117529 to 117778 (234-483bp).

CACTGTCATACTGTTTCAGAATTAAATATGCAGACCAGAAGGCTCTATACACAAGAG  
ATGGCCAGCTGCTGGTGGGAGATCCTGTTGCAGACAACTGCTGTGCAGAGAAGATC  
TGCATACTTCCTAACAGAGGCTTGGCCCGCACCAAGGTCCCCATTTTCCTGGGGATC  
CAGGGAGGGAGCCGCTGCCTGGCATGTGTGGAGACAGAAGAGGGGCCTTCCCTACA  
GCTGGAGCCATCCACCTTGCCCCCACAGGATGTGAACATTGAGGAACTGTACAAAG  
GTGGTGAAGAGGCCACACGCTTCACCTTCTTCCAGAGCAGCTCAGGCTCCGCCTTCA  
GGCTTGAGGCTGCTGCCTGGCCTGGCTGGTTCTGTGTGGCCCGGCAGAGCCCCAGC  
AGCCAGTACAGCTCACCAAGGAGAGTGAGCCCTCAGCCCGTACCAAGTTTTACTTTG  
AACAGAGCTGGTAGGGAGACAGGAACTGC

## Figure 9B

### NOVINTRA A Protein sequence:

LSYCFRIKYADQKALYTRDGQLLVGDPVADNCCAEEKICILPNRGLARTKVPIFLGIQGG  
RCLACVETEEGPSLQLEPSTLPPQDVNIEELYKGGEEATRFTFFQSSSGSAFRLEAAAWP  
GWFLCGPAEPQQPVQLTKESEPSARTKFYFEQSW

Figure 10A

**BlastN for NOVINTRA A:**

gb:GENBANK-ID:AF072476|acc:AF072476 Equus caballus interleukin-1 receptor  
antagonist secretory form (IL-1ra) gene, partial cds - Equus caballus,  
221  
bp.

Top Previous Match Next Match

Length = 221

Plus Strand HSPs:

Score = 321 (48.2 bits), Expect = 8.1e-08, P = 8.1e-08

Identities = 147/218 (67%), Positives = 147/218 (67%), Strand = Plus / Plus

Query: 253 AGGATGTGAACATTGAGGAACTGTACAA-AGGTGGTGAAGAGGCCACACGCTTCACCTTC 311  
AGG GT AACAT GA CTG CAA A G GA GAG CA CGCTTCACCTTC

Sbjct: 4 AGGCAGTTAACATCACTGACCTGAGCAAGAACAAG-GAGGAGAACAAGCGCTTCACCTTC 62

Query: 312 TTCCAGAGCAG-CTCAGGCTCCGCCTTCAGGCTT-GAGGCTGCTGCCTGGCCTGGCTGGT 369  
TCC G CA C GGC CC CC CAG CTT GAG CTGC GCCTG CCTGGCTGGT

Sbjct: 63 ATCC-GCTCAAACAGTGGCCCCACCACCAG-CTTCGAGTCTGCCGCCTGCCCTGGCTGGT 120

Query: 370 TCCTGTGTGGCCCGGCAG-AGCCCCAGCAGCCAGTACAGC-TCACCAAGGAG--AGTGA 424  
TCCT TG CG CAG AG C A C GCC GT CAGC TCACCAA AG A GA

Sbjct: 121 TCCTCTGCACGGCG-CAGGAGGCAGACCGGCCCGT-CAGCCTCACCAACAAGCCCCAAGA 178

Query: 425 GCCCT-CAGCCCGT-ACCAAGTTTACTTTGAACAGAGCTGGTAG 467  
G CCT CA GT ACCAAGTT TACTT A AG C GTAG

Sbjct: 179 GTCCTTCATG--GTCACCAAGTTCTACTTCCAGGAGGACCAGTAG 221

Figure 10B

**BlastX for NOVINTRA A:**

ptnr:TREMBLNEW-ACC:CAB59831 IL-1L1 PROTEIN - Mus musculus (Mouse), 155 aa.

Top Previous Match Next Match

Length = 155

Plus Strand HSPs:

Score = 309 (108.8 bits), Expect = 8.0e-27, P = 8.0e-27

Identities = 70/149 (46%), Positives = 93/149 (62%), Frame = +3

Query: 12 CFRIKYADQKALYTRDGQLLVGDPVADNCC-AEKICILPNRGLARTKVPIFLGIQGGSRC 188

CFR+K + K LY + QLL G A+ E+I ++PNR L + P+ LG+QGG+C

Sbjct: 8 CFRMKDSALKVLYLHNNQLLAGGLHAEKVIKGEEISVVPNRALDASLSPVILGVQGGSC 67

Query: 189 LACVETEEGPSLQLEPSTLPPQDVNIEELYKGGEETRFTFFQSSSGSAFRLEAAAWPGW 368

L+C TE+GP L+LEP VNI ELY G +E+ FTF++ G E+AA+PGW

Sbjct: 68 LSC-GTEKGPILKLEP-----VNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGW 119

Query: 369 FLCGPAEPQQPVQLTKESEPSAR----TKFYFEQ 458

FLC E QPV+LT+ E A T FYF+Q

Sbjct: 120 FLCTSPEADQPVRLTQIPEDPAWDAPITDFYFQQ 153

>ptnr:SPTREMBL-ACC:Q9UBH0 FIL1 DELTA (INTERLEUKIN-1 LIKE PROTEIN 1)

(INTERLEUKIN-1 RECEPTOR ANTAGONIST HOMOLOG 1) (INTERLEUKIN-1 DELTA) - Homo sapiens (Human), 155 aa.

Plus Strand HSPs:

Score = 285 (100.3 bits), Expect = 4.3e-24, P = 4.3e-24

Identities = 67/149 (44%), Positives = 89/149 (59%), Frame = +3

Query: 12 CFRIKYADQKALYTRDGQLLVGDPVADNCC-AEKICILPNRGLARTKVPIFLGIQGGSRC 188

CFR+K + K LY + QLL G A E+I ++PNR L + P+ LG+QGG+C

Sbjct: 8 CFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEEISVVPNRWLDASLSPVILGVQGGSC 67

Query: 189 LACVETEEGPSLQLEPSTLPPQDVNIEELYKGGEETRFTFFQSSSGSAFRLEAAAWPGW 368

L+C +E P+L LEP VNI ELY G +E+ FTF++ G E+AA+PGW

Sbjct: 68 LSCGVGQE-PTLTLEP-----VNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGW 119

Query: 369 FLCGPAEPQQPVQLTKESEPSAR----TKFYFEQ 458

FLC E QPV+LT+ E T FYF+Q

Sbjct: 120 FLCTVPEADQPVRLTQLPENGGWNAPITDFYFQQ 153

Figure 11

### ClustalW for NOVINTRA A:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g. L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

### Sequences used:

- 1) X86458spoil1 (cDNA encoding a human spoil-I protein); X86459spoil2 (cDNA encoding a human spoil-II protein); X89432-il1delta (human IL-1 delta encoding cDNA); Q14628 (intracellular IL-1 receptor antagonist type II-human); G512010il1beta (ovine IL-1 beta - pig)

```

X86458spoil1  - - - - - DNHTMRGTPGDA - - - - - DGG - - GRAVYQS - - - - -
X86459spoil2  - - - - - DNHTMRGTPGDA - - - - - DGG - - GRAVYQSSESNAVGMGLWRLRPSALTLS PV
X89432-il1delta - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Q14628        - - - - - MALADLYEEG - - - - - GGG - - GGEDNADSK - - - - - E - - - - -
il1beta-1DNA  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
G512010il1beta MATVPEP I NEVMAYYSD ENELLFEVDGP KQMK SCTQHLDLGSMGDGNIQLQISHQLYNKS

X86458spoil1  - - - - - - - - - - - MCKPITGT INDLNQVWTLQGONLVAVPRSDS - - MTPVTVAVIT
X86459spoil2  EAPAFSAPCTLPFPVPYCKPITGT INDLNQVWTLQGONLVAVPRSDS - - MTPVTVAVIT
X89432-il1delta - - - - - - - - - - - MVL SGALCFRMKDSALKVLYLHNNQLLAGGLHA - - - - - GKVIKGEET
Q14628        - - - - - TCRPSGRKSSKMQAFRIWDNNQKTFYLRNNQLVAGYLQC - - - - - PNWNLEEKI
il1beta-1DNA  - - - - - - - - - - - GPSALSYCFRIKYADOKALYTRDGLVGDPA - - - - - DNCCAELI
G512010il1beta FRQVVSVIYAMEKLRSRAYEHVFRDDDLRSISSTIFE EEPVIFETSSDELLCDAAVQSVK

X86458spoil1  CKYPEAL - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
X86459spoil2  CKYPEAL - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
X89432-il1delta SVMFNRW - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Q14628        DVVP - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
il1beta-1DNA  CIPNRRG - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
G512010il1beta CKIQDREQKSLVLDSPCVLKALHLLSQEMSREVVFCMSFVQGEERDNKIPVALGIRDKNL

X86458spoil1  CLYCEKVGEQPTLQLKE - - - - - QKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPD
X86459spoil2  CLYCEKVGEQPTLQLKE - - - - - QKIMDLYGQPEPVKPFIFYRAKTGRTSTLESVAFPD
X89432-il1delta CLSCG-VCQEPTLTLEP - - - - - VNIMELYLGAKESKSFTFYRRDMGLTSSFESAAPG
Q14628        CLSCVKSGETRLQLEA - - - - - VNITDLSNRKQDKRFAFIRSDSGPTTSFESAACPG
il1beta-1DNA  CLACVETEEGFSLQLEPSTLPPQDVNI EELYKGGEEATRFTTFQSSSGSAFRLEAAAPG
G512010il1beta YLSCVKKCDTPTLQLEE - - - - - VDPKMYPKRNMEKRFVFKTEIKNTVEFESVLYPN

X86458spoil1  WFLASS - KRDQP ILTSELGKS - - - - YNTAFELNIND -
X86459spoil2  WFLASS - KRDQP ILTSELGKS - - - - YNTAFELNIND -
X89432-il1delta WFLCTVPEADQPVRLTQLPENGGWNAPI TDFYFQCCD -
Q14628        WFLCTAMEADQPVSLTNMPDEG - - - - VMYTKFYFQED -
il1beta-1DNA  WFLCGPAEPQQPVQLTKSEPS - - - - ARTKFYFEQSW -
G512010il1beta WYLSISQIEKPVFLGRFRGGQ - - - - DITDFRMETLSP

```

## Figure 12A

### NOVINTRA B cDNA sequence:

Regions 116968 to 117028 (1-61bp), 104067 to 104164 (62-159bp), 101958 to 102098 (160-300bp), and 100915 to 101134 (301-520bp) of genomic clone (-frame).

ATGGGCACACCTGGCCTGGCCCTGCATGCAGACTGGACGGTGAGCCAGGACTTCTG  
CAGGACACCCAAATCCTATGCTATTCGTGATTCTCGACAGATGGTGTGGGTCCTGAG  
TGGAAATTCTTTAATAGCAGCTCCTCTTAGCCGCAGCATTAAAGCCTGTCACTCTTCAT  
TTAATAGCCTGTAGAGACACAGAATTCAGTGACAAGGAAAAGGGTAATATGGTTTA  
CCTGGGAATCAAGGGAAAAGATCTCTGTCTCTTCTGTGCAGAAATTCAGGGCAAGCC  
TACTTTGCAGCTTAAGGAAAAAATATCATGGACCTGTATGTGGAGAAGAAAGCAC  
AGAAGCCCTTTCTCTTTTCCACAATAAAGAAGGCTCCACTTCTGTCTTTCAGTCAGT  
CTCTTACCCTGGCTGGTTCATAGCCACCTCCACCACATCAGGACAGCCCATCTTTCTC  
ACCAAGGAGAGAGGCATAACTAATAACACTAACTTCTACTTAGATTCTGTGGAATA  
AATCCAGC

## Figure 12B

### NOVINTRA B Protein sequence:

MGTPGLALHADWTVSQDFCRTPKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLI  
ACRDTEFS DKEKGNMVYLG IKGKDLCLFCAEIQGKPTLQLKEKNIMDL YVEKKAQKPFL  
FFHNKEGSTSVFQSVSYPGWFIATSTTSGQPIFLTKEGITNNTNFYLD SVE



Figure 13A

**BlastN for NOVINTRA B:**

gb:GENBANK-ID:PIGIRAP|acc:L38849 Sus scrofa (clone p0328w) IRAP1 mRNA,  
complete cds - Sus scrofa, 563 bp (RNA).

Top Previous Match Next Match

Length = 563

Plus Strand HSPs:

Score = 309 (46.4 bits), Expect = 1.1e-07, P = 1.1e-07

Identities = 197/331 (59%), Positives = 197/331 (59%), Strand = Plus / Plus

Query: 169 CATTTAATAGCCTGTA-GAGAC-ACAGAATTCAGTGACAAGGAAAAGGGTAATATGGTTT 226  
CA TA TA CTG A GAGA A AGA T GTG C G AG T AT T GT T  
Sbjct: 217 CAAATACTAACTGGAAGAGAAGATAGATGTG-GTGCCT--GTTGAGCCTCATTTTGTGT 273

Query: 227 ACCTGGGAATCAAGGGAAAAGATCTCTGTCTCTTCTGTGCAGAAATTCAGGGCAAGCC-T 285  
CCTGGG ATC A GGA A CT TG CT T CTGTG AA T C GG A G T  
Sbjct: 274 TCCTGGGGATCCATGGAGGGAAGCTGTGCCTGTCTGTGTC-AAGT-CTGGTGATGAGAT 331

Query: 286 -ACTTTGCAGCTTAAGGAAAAAATATCATGGACCTGTATGTGGAGAAGAAAGCAC-AGA 343  
A T CAG T A G A AA ATCA GACCTG A GA AG AGCA A A  
Sbjct: 332 GAAGCTCCAGTTGGACGCGAGTTAACATCACAGACCTG-AGAAAGAACAGCGAGCAGGACA 390

Query: 344 AGCCCTTTTCTCTTTTCCACAATAA-AGAAGGCTCCACTTCTGTCTTTTCAGTCAGTCTCT 402  
AGC CTT CTT TCC C A AG GGC CCAC C CTTT AGTCAG C C  
Sbjct: 391 AGCGCTTCACCTTCATCCGCTCCGACAGT-GGCCCCACCACCAGCTTTGAGTCAGCCGCC 449

Query: 403 TACCCTGGCTGGTTCATAGCCACCTCCACCACAT-CAGGACAGCCCATCTTTCTCACCAA 461  
T CCTGGCTGGTTC T CAC T CAC A A CAG CAGCC T CTCACCAA  
Sbjct: 450 TGTCTGGCTGGTTCCTCTGCAC-TGCACTAGAGGCAGACCAGCCTGTTGGCCTCACCAA 508

Query: 462 GGAGAG-AG-GCATAACT-AATAACACTAACTTCTACTT 497  
G A GCA T AA CAC AA TTCTACTT  
Sbjct: 509 CACGCCCAAAGCAGCCGTCAGGTCACCAAGTTCTACTT 547

Figure 13B

**BlastX for NOVINTRA B:**

>ptnr:SPTREMBL-ACC:Q9UHA5 FIL1 ETA - Homo sapiens (Human), 157 aa.  
Plus Strand HSPs:

Score = 777 (273.5 bits), Expect = 3.2e-76, P = 3.2e-76  
Identities = 149/149 (100%), Positives = 149/149 (100%), Frame = +1

Query: 64 PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLLIACRDTEFSDKEKGNMVYLGIKG 243  
PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLLIACRDTEFSDKEKGNMVYLGIKG  
Sbjct: 9 PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLLIACRDTEFSDKEKGNMVYLGIKG 68

Query: 244 KDLCLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA 423  
KDLCLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA  
Sbjct: 69 KDLCLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIA 128

Query: 424 TSTTSGQPIFLTKERGITNNTNFYLDSE 510  
TSTTSGQPIFLTKERGITNNTNFYLDSE  
Sbjct: 129 TSTTSGQPIFLTKERGITNNTNFYLDSE 157

>ptnr:SPTREMBL-ACC:Q9NZH7 INTERLEUKIN-1 HOMOLOG 2 - Homo sapiens (Human), 164 aa.  
Plus Strand HSPs:

Score = 411 (144.7 bits), Expect = 1.9e-37, P = 1.9e-37  
Identities = 80/85 (94%), Positives = 81/85 (95%), Frame = +1

Query: 64 PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLLIACRDTEFSDKEKGNMVYLGIKG 243  
PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLLIACRDTEFSDKEKGNMVYLGIKG  
Sbjct: 9 PKSYAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLLIACRDTEFSDKEKGNMVYLGIKG 68

Query: 244 KDLCLFCAEIQGKPTLQLKEKNIMD 318  
KDLCLFCAEIQGKPTLQLK + D  
Sbjct: 69 KDLCLFCAEIQGKPTLQLKLQGSQD 93

>ptnr:TREMBLNEW-ACC:AAF02757 INTERLEUKIN-1 RECEPTOR ANTAGONIST HOMOLOG  
- Homo sapiens (Human), 155 aa.

Top Previous Match Next Match  
Plus Strand HSPs:

Score = 197 (69.3 bits), Expect = 5.9e-15, P = 5.9e-15  
Identities = 52/147 (35%), Positives = 76/147 (51%), Frame = +1

Query: 73 YAIRDSRQMVWVLSGNSLIAAPLSRSIKPVTLHLLIACRDTEFSDKEKGNMVYLGIKGKDL 252  
+ ++DS V L N L+A L K + I+ + D ++ LG++G  
Sbjct: 9 FRMKDSALKVLYLHNNQLLAGGLHAG-KVIKGEEISVVPNRWLDASLPVI-LGVQGSQ 66

Query: 253 CLFCAEIQGKPTLQLKEKNIMDLYVEKKAQKPFLFFHNKEGSTSVFQSVSYPGWFIATST 432  
CL C Q +PTL L+ NIM+LY+ K K F F+ G TS F+S +YPGW+ T  
Sbjct: 67 CLSCGVGQ-EPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVP 125

Query: 433 TSGQPIFLTK--ERGITNN--TNFYLDSE 510  
+ QP+ LT+ E G N T+FY +  
Sbjct: 126 EADQPVRILTQLPENGGWNAPITDFYFQQCD 155

Figure 14

# ClustalW for NOVINTRA B:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (*i.e.*, regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (*e.g.* L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

## Sequences used:

- 2) X86459spoil2 (cDNA encoding a human spoil-II protein; Q14628 (intracellular IL-1 receptor antagonist type II-human); AAF02757 (intracellular IL-1 receptor antagonist type II-human); G512010; (OVINE IL-1 BETA - Pig)

## Multiple Alignment:

```

il1beta2DNA      - - - - - M G T P G L A L H A D W T V S Q D - - - -
X86459spoil2     - - - - - D N H T M R G T P C D A D G G R A V Y Q S S E S N
Q14628           - - - - - M A L A D T Y E E G G C G G G E G E D N A D S K E -
AAF02757         - - - - -
G512010il1beta  M A T V P E P I N E V M A Y S D E N E L L F E V D G P K Q M K S C T Q H L D L G S M C D G N I Q L Q I S H Q L Y N K S

il1beta2DNA      - - - - - F C R T - - - - - P K S Y A I R - - - - -
X86459spoil2     A V G M G L W R L R P S A L T L S P V E A P A F S A P L C T L P F P P V C K - P T G T I N - - - - -
Q14628           - - - - - T I C R P S G R K S S K - M Q A F R I W - - - - -
AAF02757         - - - - - M V L S G - A L C F R M K - - - - -
G512010il1beta  F R Q V V S V I V A M E K L R S R A Y E H V F R D D D L R S I L S F I F E E E P V I F E T S S D E L L C D A A V Q S V K

il1beta2DNA      - - - - D S R C M V W V L S G N S L I A A P L S R S I K P V T L H L I A C R D T E F S D K - E K G N M V Y L G I K E K D
X86459spoil2     - - - - D L N Q Q V W T L Q G Q N L V A V P R S D S V T P V T V A V I T C K Y P E A L E Q - G R G D P I V L G I Q N P E
Q14628           - - - - D V N Q K T T Y L R N N Q L V A G Y L Q - - G P N V N L E E K I D V V P - - - - - I E P H A L F L G I H G G K
AAF02757         - - - - D S A L K V L Y L H N N Q L L A C G L H - - A G K V I K G E E I S V V P N R W L D - A S L S P V I L G V Q E G S
G512010il1beta  C K L Q D R E Q K S L V L D S - P C V L K A L H L L S Q E M S R E V V F C M S F V Q G E E R D N K I P V A L G I R D K N

il1beta2DNA      L C L F C A E I Q G K P T L Q L K E K N I M D L Y V E K K A Q K P F L F F H N K E G S T S V F Q S V S Y P G W F I A T S
X86459spoil2     M C L Y C E K V G E Q P T L Q L K E Q K I M D L Y G Q P E P V K P F L F Y R A K T G R T S T L E S V A F P D W F I A S S
Q14628           M C L S C V K S G D E T R L Q L E A V N I T D L S E N R K Q D K R F A F I R S D S G P T T S F E S A A C P G W F L C T A
AAF02757         Q C L S C G - V G Q E P T L T L E P V N I M E L Y L G A K E S K S F T F Y R R D M G L T S S F E S A A Y P G W F C T V
G512010il1beta  L Y L S C V K K G D T P T L Q L E E V D P K V Y P - K R N M E K R F V F Y K T E I K N T V E F E S V L Y P N W Y I S T S

il1beta2DNA      T T S G Q P I F L T K E R - - G - - I T N N T N F Y L D S V E -
X86459spoil2     K R D - Q P I I L T S E L - - G - - K S Y N T A F E L N I N D -
Q14628           M E A D Q P V S L T N M P D E G - - V M V T K F Y F Q E D E -
AAF02757         P E A D Q P V R L T Q L P E N G G W N A P I T D F Y F Q Q C D -
G512010il1beta  Q I E E K P V F L G R F R - - G - - G Q D I T D F R M E T L S P

```

**Figure 15A**

**NOVINTRA C cDNA sequence:**

Region 168700 to 168880 (1-181bp), 169967 to 170118 (182-331bp) and 170635 to 170694 (332-391bp) of the genomic clone (- frame).

GATATCAATCATCGGGTGTGGGTTCCTTCAGGACCAGACGCTCATAGCAGTCCCGAGG  
AAGGTGTTCCCAAGTCACTATTGCCTTAATCTCATGCCGACATGTGGAGACCCCTTGAG  
AAAGACAGAGGGAACCCCATCTACCTGGGCCTGAATGGACTCAATCTCTGCCTGAT  
GTGTGCTAAAGTCGGGGACCAGCCACACTGCAGCTGAAGCTTCAGGAAAAGGATA  
TAATGGATTTGTACAACCAACCCGAGCCTGTGAAGTCCTTTCTCTTCTACCACAGCC  
AGAGTGGCAGGAACCTCCACCTTCGAGTCTGTGGCTTTCCTGGCTGGTTCATCGCTG  
TCAGCTCTGAAGGAGGCTGTCCTCTCATCCTTACCCAAGAACTGGGGAAAG

**Figure 15B**

**NOVINTRA C Protein sequence:**

DINHRVWVLQDQTLIAVPRKVFPVTIALISCRHVETLEKDRGNPIYLGLNGLNLCLMCAK  
VGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSEGGC  
PLILTQELGK

Figure 16A

**BlastN for NOVINTRA C:**

gb:GENBANK-ID:PIGIRAP|acc:L38849 Sus scrofa (clone p0328w) IRAP1 mRNA,  
complete cds - Sus scrofa, 563 bp (RNA).

Top Previous Match Next Match

Length = 563

Minus Strand HSPs:

Score = 274 (41.1 bits), Expect = 5.6e-06, P = 5.6e-06

Identities = 150/244 (61%), Positives = 150/244 (61%), Strand = Minus / Plus

Query: 258 TCTACCTGGGCCTGAATGGACTCAATCTCTGCCTGATGTGTGCTAAAGTCGGG-GACCAG 200  
T T CCTGGG T ATGGA AA CT TGCCTG TGTG T AAGTC GG GA AG  
Sbjct: 271 TGTTCTGGGGATCCATGGAGGGAAGCTGTGCCTGTCCTGTG-TCAAGTCTGGTGATGAG 329

Query: 199 CCCACACTGCAGCTGAAGCTTCAGGAAAAGGATATAATGGATTTGTACAACCAACC-CGA 141  
A CT CAG TG A C CAG AA AT A GA TG A AA AAC CGA  
Sbjct: 330 ATGAAGCTCCAGTTGGA-CG-CAGTTAAC-----ATCACAGACCTG-AGAAAGAACAGCGA 382

Query: 140 GCCTGTGAAGTCCTTTCTCTTCTACCACAGCCAGAGTGGCAGGAACTCCACCTTCGAGTC 81  
GC G AAG CTT CTTC CC C C A AGTGGC A C CCA CTT GAGTC  
Sbjct: 383 GCAGGACAAGCGCTTCACCTTCATCCGCTCCGACAGTGGCCCCACCACCAGCTTTGAGTC 442

Query: 80 TGTGGCTTTCCCTGGCTGGTTCATCGCTGT-CAGCTCTGAAGG-AGGCTGTCCTCTCATC 23  
G GC T CCTGGCTGGTTC TC TG C GC CT AGG AG C CCT T C  
Sbjct: 443 AGCCGCCTGTCTGGCTGGTTCCTC--TGCACTGCACTAGAGGCAGACCAGCCTGTTGGC 500

Query: 22 CTTACCCA 15  
CT ACC A  
Sbjct: 501 CTCACCAA 508

Figure 16B

**BlastX for NOVINTRA C:**

>ptnr:SPTREMBL-ACC:Q9UHA7 FIL1 EPSILON - Homo sapiens (Human), 158 aa.

Plus Strand HSPs:

Score = 646 (227.4 bits), Expect = 2.4e-62, P = 2.4e-62  
Identities = 126/130 (96%), Positives = 127/130 (97%), Frame = +1

Query: 1 DINHRVWVLQDQTLIAVPRK--VFPVTIALISCRHVETLEKDRGNPIYLGGLNGLNLCCLMC 174  
DINHRVWVLQDQTLIAVPRK + PVTIALISCRHVETLEKDRGNPIYLGGLNGLNLCCLMC  
Sbjct: 17 DINHRVWVLQDQTLIAVPRKDRMSPVTIALISCRHVETLEKDRGNPIYLGGLNGLNLCCLMC 76

Query: 175 AKVGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSEG 354  
AKVGDQPTLQLK EKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSEG  
Sbjct: 77 AKVGDQPTLQLK--EKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSEG 134

Query: 355 GCPLILTQELGK 390  
GCPLILTQELGK  
Sbjct: 135 GCPLILTQELGK 146

>ptnr:SPTREMBL-ACC:Q9NZH8 INTERLEUKIN-1 HOMOLOG 1 - Homo sapiens, 169 aa.

Plus Strand HSPs:

Score = 416 (146.4 bits), Expect = 5.7e-38, P = 5.7e-38  
Identities = 82/130 (63%), Positives = 101/130 (77%), Frame = +1

Query: 1 DINHRVWVLQDQTLIAVPRK--VFPVTIALISCRHVETLEKDRGNPIYLGGLNGLNLCCLMC 174  
D+N +VW LQ Q L+AVPR V PVT+A+I+C++ E LE+ RG+PIYLG+ +CL C  
Sbjct: 29 DLNQVWTLQGGQNLVAVPRSDSVTPVTVAVITCKYPEALEQGRGDPIYLGGLNGLNLCCLMC 88

Query: 175 AKVGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFESVAFPGWFIAVSSEG 354  
KVG+QPTLQLK E+ IMDLY QPEPVK FLFY +++GR ST ESVAFP WFIA SS+  
Sbjct: 89 EKVGEQPTLQLK--EQKIMDLYGQPEPVKPLFYRAKTGRTSTLESVAFPDWFIA-SSKR 145

Query: 355 GCPLILTQELGK 390  
P+ILT ELGK  
Sbjct: 146 DQPIILTSELGK 157

>ptnr:SWISSPROT-ACC:P25085 INTERLEUKIN-1 RECEPTOR ANTAGONIST PROTEIN  
PRECURSOR (IL-1RA) (IL-1RN) (IRAP) - Mus musculus (Mouse), 178 aa.

Top Previous Match Next Match

Minus Strand HSPs:

Score = 173 (60.9 bits), Expect = 2.1e-12, P = 2.1e-12  
Identities = 35/81 (43%), Positives = 50/81 (61%), Frame = -1

Query: 259 IYLGGLNGLNLCCLMCAKVGDQPTLQLKLQEKDIMDLYNQPEPVKSFLFYHSQSGRNSTFES 80  
++LG++G LCL CAK GD ++L+L+E +I DL E K F F S+ G ++FES  
Sbjct: 82 VFLGIHGGKCLSCAKSGDD--IKLQLEEVNITDLSKNKEEDKRFTFIRSEKGPPTS FES 139

Query: 79 VAFPGWFIAVSSEGGCPLILT 17  
A PGWF+ + E P+ LT  
Sbjct: 140 AACPGWFLCTTLEADRPVSLT 160

Figure 17

# ClustalW for NOVINTRA C:

Based on this alignment, black outlined amino acid residues indicate regions of conserved sequence (i.e., regions that may be required to preserve structural or functional properties); greyed amino acid residues can be mutated to a residue with comparable steric and/or chemical properties without altering protein structure or function (e.g. L to V, I, or M); non-highlighted amino acid residues can potentially be mutated to a much broader extent without altering structure or function.

## Sequences used:

- 1) IL1X\_MOUSE: Locus - IL1X\_MOUSE; Accession - P25085; PID - g124313
- 2) IL1X\_HORSE: Locus - IL1X\_HORSE; Accession - O18999; PID - g6166230
- 3) AAF02757\_HUMAN: Locus - AF186094\_1; Accession - AAF02757; PID - g6049805

```

IL1X_MOUSE      MEICWGPYSHLISLLFLFHSEACRPSSGKRPCCKMQAFRIWDTNQKTFYLRNNQLIAGY
IL1X_HORSE      MEIRRRSVRHLIS-LLFLFYSETACHPLGKRPCCKMQAFRIWDVNQKTFYLRNNQLVAGY
AAF02757_Human  -----MVLSGALC-----FRMKDSALKVLYLHNNQLIAGG
Novel_IL1       -----DINHRVWVLQDQTLIAVP

IL1X_MOUSE      LGGPNIKLE---EKIDVVP-----IDLHSVFLGHHGGKLCCLSCAKSGDDIKLQLEEVN-
IL1X_HORSE      LQESNTKLQ---EKIDVVP-----IEPDALFLGLHGRKLCLACVKSGDEIRFQLEAVN-
AAF02757_Human  LHAGKVIKGG---EEISVYPNRWLDASLSPVILGVQGGSCCLSCG-VGQEP TLTLEPVN-
Novel_IL1       RKVFPMTIALISCRHYETLEK---DRGNPILGLNGLNLCLMCAKVG DQPTLQLKLQEK

IL1X_MOUSE      -ITDLSKNKEEDKRFTFIRSEKGP TTSFESAACPGWFLCTTLEADRPVSLTNTPEE---P
IL1X_HORSE      -ITDLSKNKEENKRFTFIRSNSGPTTSFESAACPGWFLCTAQEADRPVSLTKPKE---S
AAF02757_Human  -IMELYLGAKESKSFTFYRRDMGLTSSSFESAAYPGWFLCTVPEADQPVRLTLPENGGWN
Novel_IL1       DIMDLYNQPEPVKSELFYHSOSGRNSTFESMAFPGWFLAVSSEGGCPILITQELGK---

IL1X_MOUSE      LITVTKFYFQEDQ
IL1X_HORSE      FMVTKFYLLQEDQ
AAF02757_Human  APITDFYFQQCD
Novel_IL1       -----

```

Figure 18

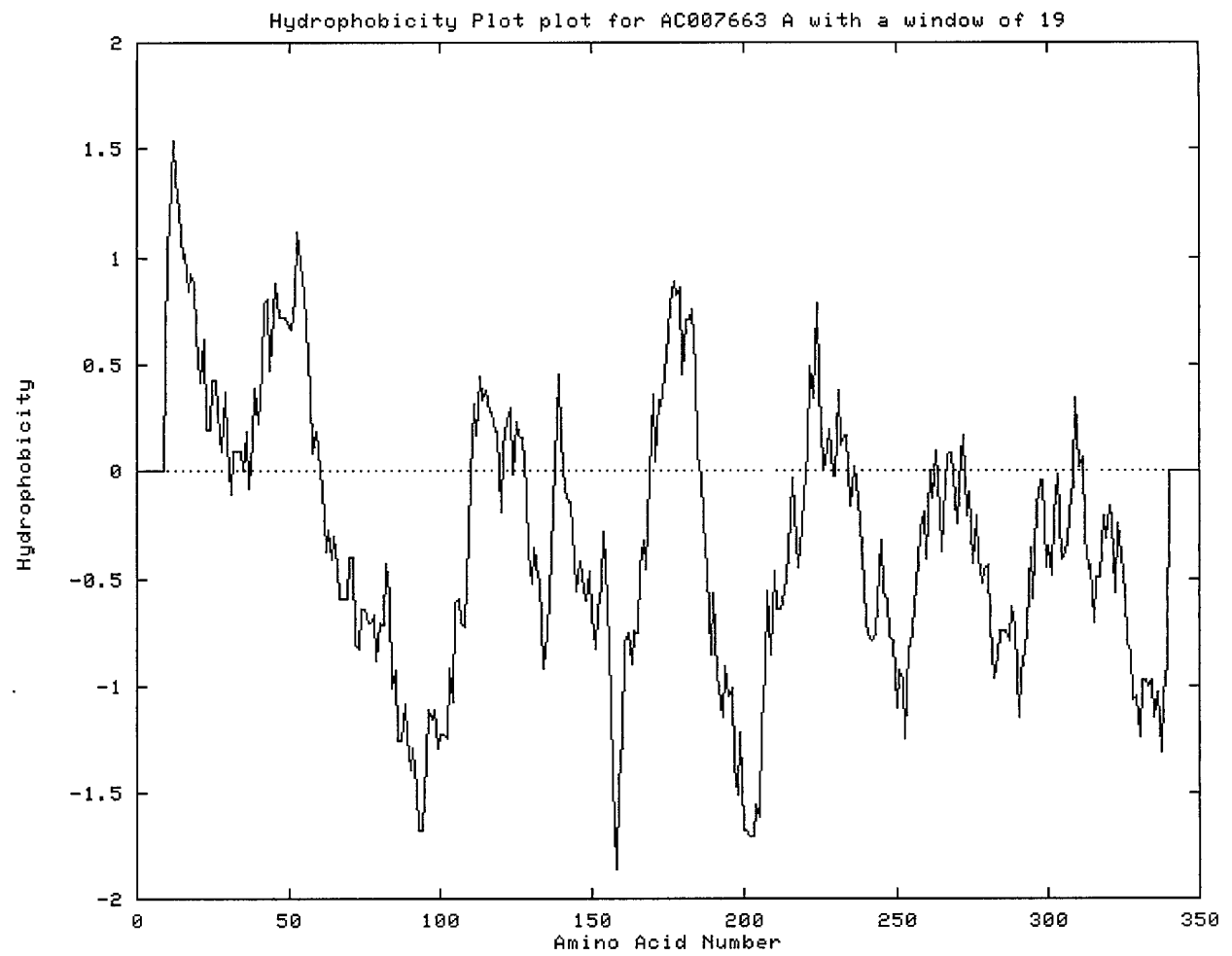
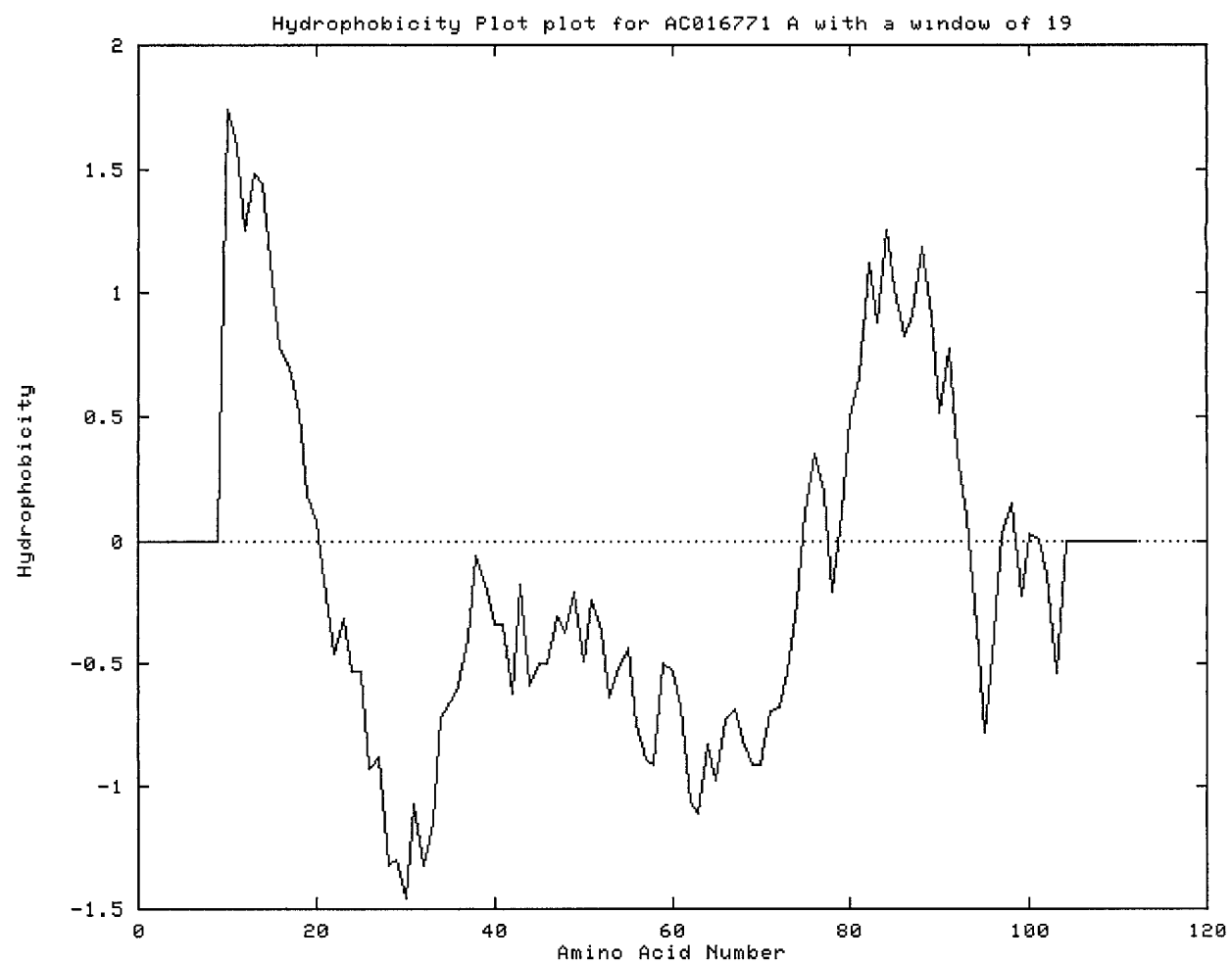




Figure 19



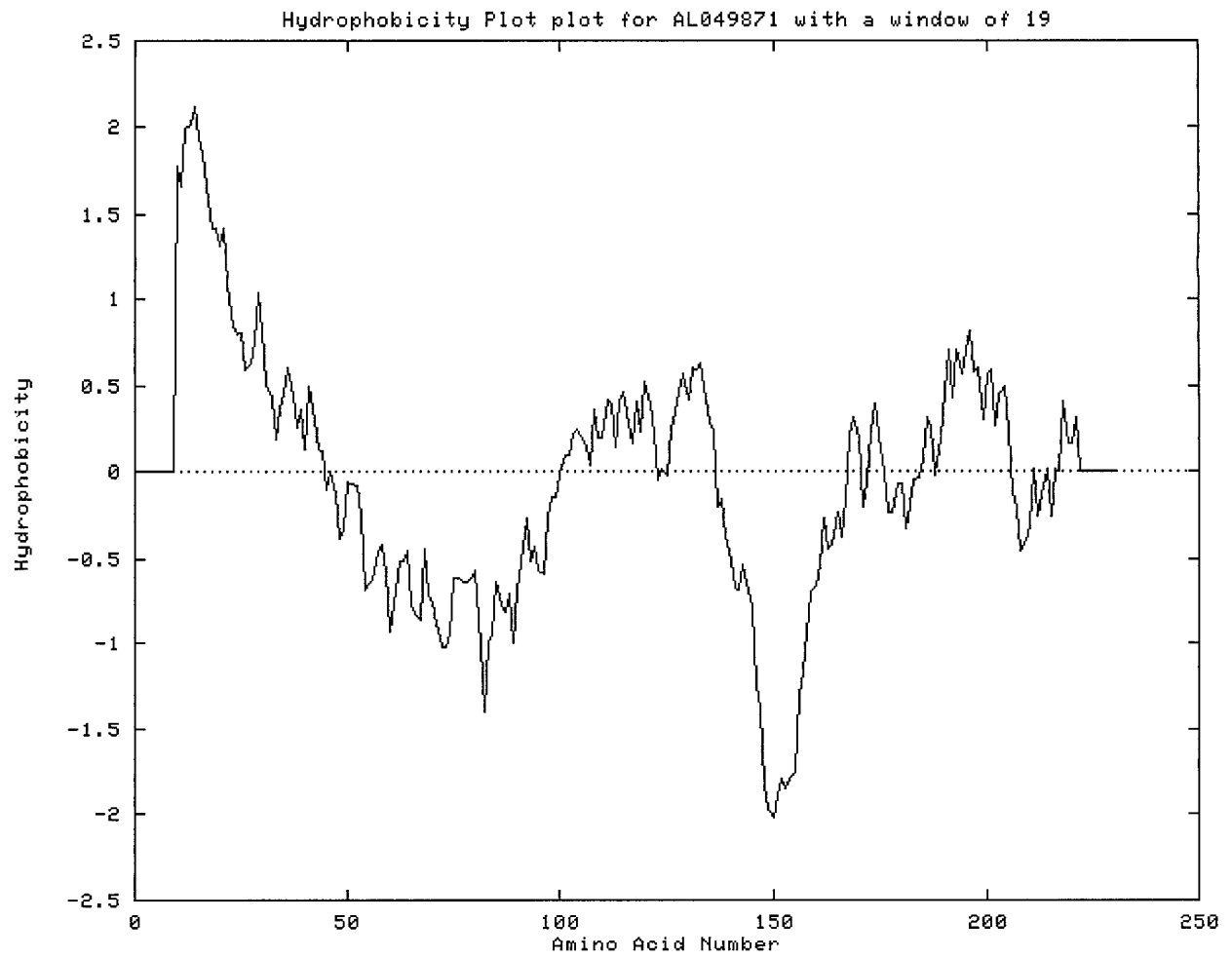
[illegible]

Figure 21

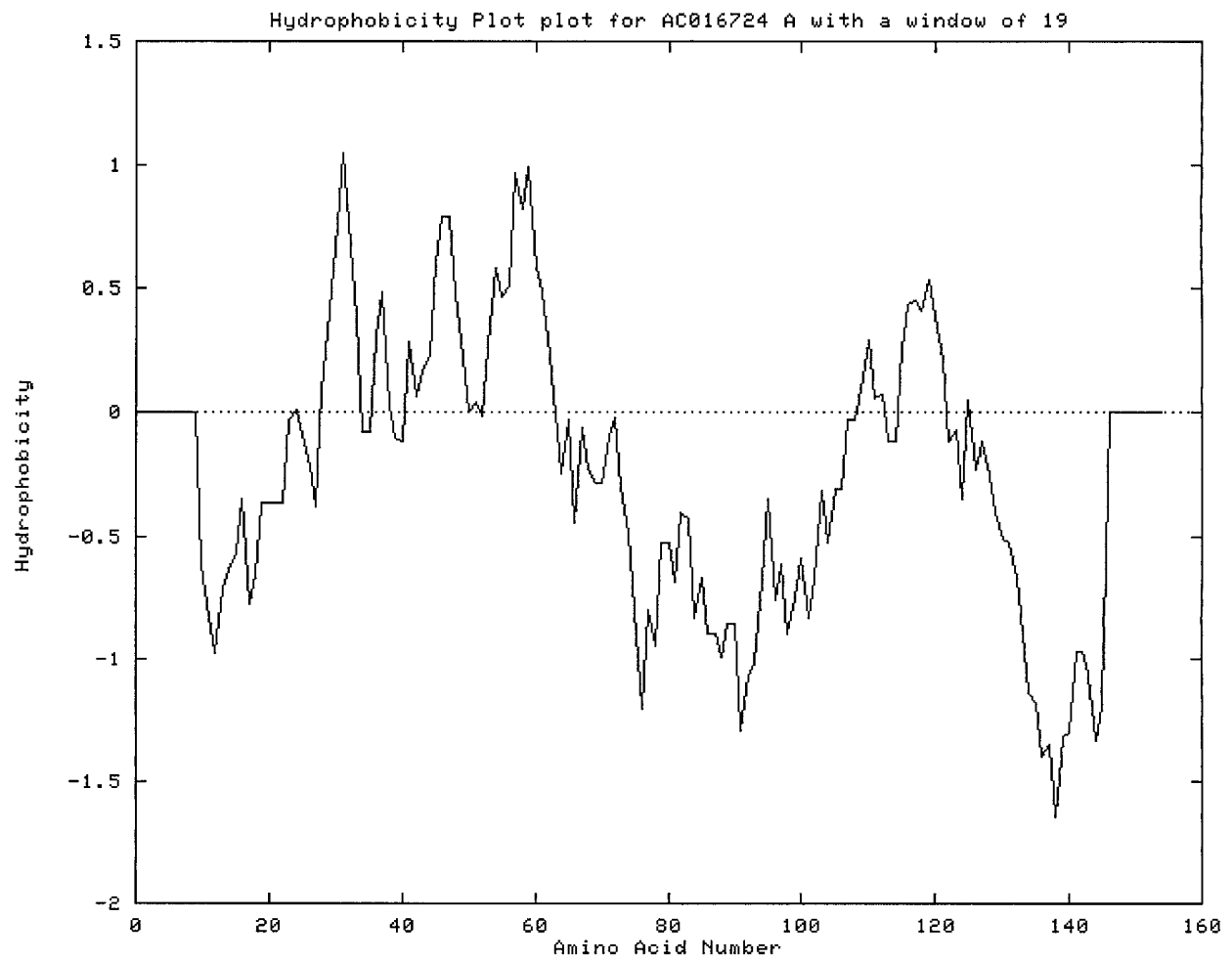
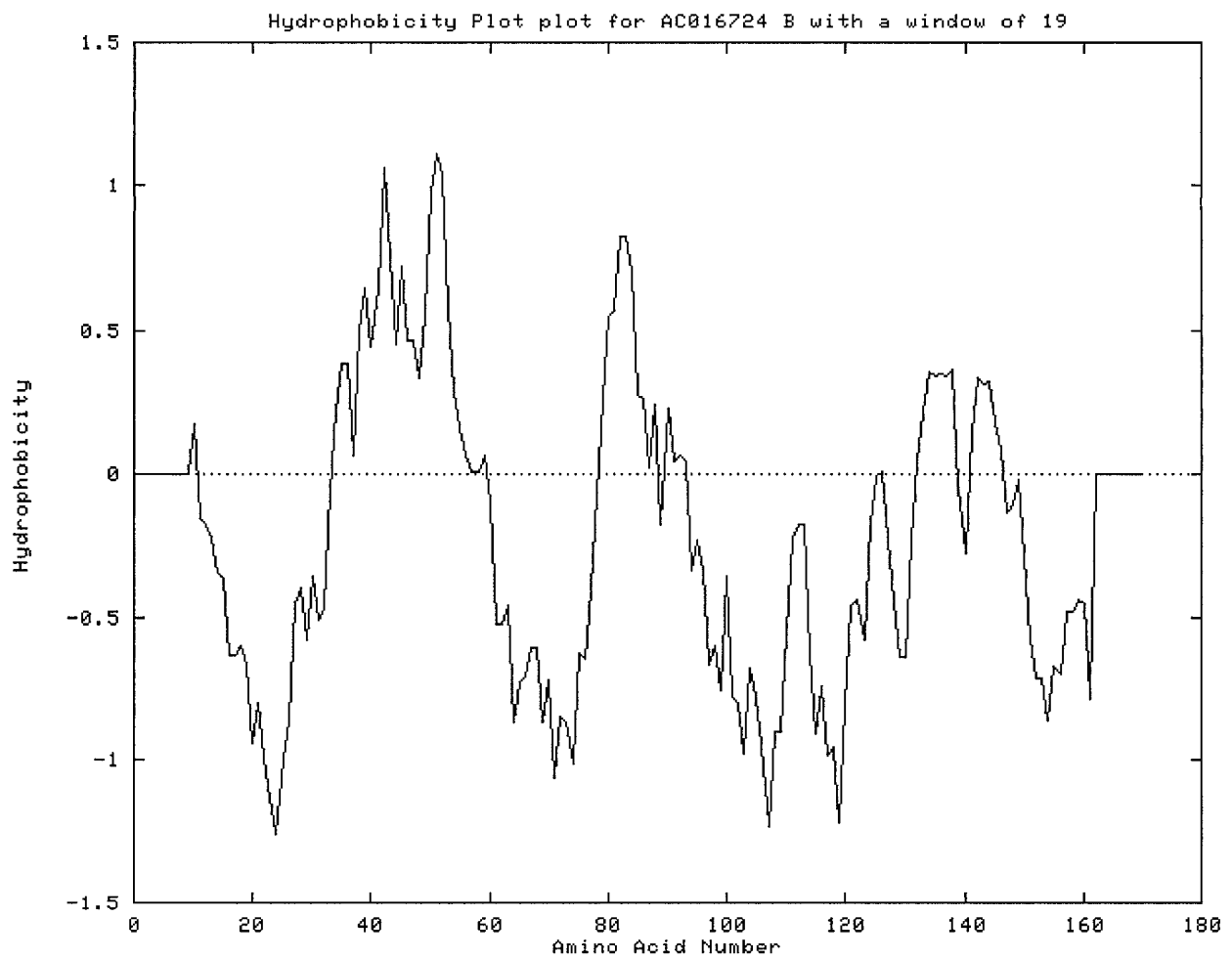


Figure 22



AC016724 B

Figure 23

